



AI Policy

Updated: 16 June 26

Policy Scope

This policy applies to all trustees, other volunteers, employees, contractors, and third-party representatives working on behalf of FEAST With Us (hereby known as FEAST). Its requirements should be reflected in other policies and procedures, agreements and contracts, as necessary.

Definition

FEAST defines Artificial Intelligence (AI) as the ability of machines or software to perform tasks that would normally require human intelligence. AI systems can process data, learn from it, and make decisions or predictions based on that data. AI is a broad field that encompasses many different types of systems and approaches to machine intelligence, including rule-based AI, machine learning, neural networks, natural language processing and robotics.

We refer to 'Created Content' throughout this policy. FEAST defines Created Content, in regard to this AI Policy, as content created entirely using prompts within a GenAI system. We would not consider text that has been condensed, reframed or reviewed to correct grammar as Created Content, and as such would not be obliged to disclose the use of AI as part of that editing process.

[See Appendix 1 for a glossary of terms.](#)

All key AI decisions and proposals will be subject to scrutiny and approval by the FEAST Trustee Board. They will be advised on any concerns or breaches in AI use and will review this policy and FEAST's AI performance annually to keep up with evolving AI technologies and ethical standards.

Use of AI by FEAST will have appropriate human oversight with humans being responsible for making all final decisions on their output. FEAST will maintain oversight by monitoring AI systems' performance, impact, and compliance with this policy on an ongoing basis.

To support this, FEAST will create any necessary guidelines on the collection, use and storage of data. This will ensure accountability for the decisions made by AI systems, which may include measures such as auditing, reporting and review processes and the use of algorithms in decision-making, including the steps FEAST will take to ensure these are as fair and unbiased as reasonably possible.

Policy on Management of AI

FEAST will support our people in adapting to the changes AI will bring, by providing them with appropriate support and skills development and considering their needs, when designing roles and work procedures. The requirements of FEAST's AI policy will be embedded in other relevant policies and procedures, contracts, agreements and other documentation, such as job descriptions, where necessary. FEAST will ensure that those at FEAST with responsibilities for or involvement in AI, understand FEAST's AI policy, their responsibilities in delivering this and are accountable for doing so.

AI Risk Management

FEAST's AI risk analysis has included any specific groups who may be at risk and other reasonably foreseeable uses of the technology, including accidental or malicious misuse. The risks have been identified and quantified, and the avoidance and mitigation action put in place will ensure that the level of risk remains within acceptable limits. These risks include, but are not limited to:

Biases: AI sweeps the internet for information, but the results are not always presented as a neutral viewpoint.

Plagiarism: When using AI, the human creators are often unaware of what the sources are, so they could be using the work of others without proper citations or acknowledgement.

Security: AI tools, alongside the information and data sets used as inputs, could be targeted by hackers.

AI Data Protection & Privacy

Insofar as reasonably possible, FEAST will:

- Use accurate, fair, and representative data sets to ensure these are inclusive.
- Not include personal identification or personal characteristics data, or at least pseudo-anonymise or de-identify it.
- Ensure FEAST's data consent procedures are always simple and clear and obtain user consent when using AI systems that process personal data.
- Reflect FEAST's use of AI in our privacy statement to ensure users know when their data is being used by AI, whether AI is making decisions about them and, if so, what these decisions are.

FEAST is aware of the ICO guidance on AI and data protection and have reflected any additional requirements in our policies and procedures.

AI Ethics

FEAST is committed to genuinely engaging with our stakeholders to ensure that our AI is aligned with their needs and values. We factor into our risk analysis, any exclusion or detriment to them based on their identity. FEAST will take reasonable steps to avoid or minimise any exclusion or detriment and transparently communicate this.

FEAST will only ever use AI image generation to create design assets, and not imagery depicting real-life settings.

FEAST will make our AI systems and content as accessible as possible. Insofar as reasonably possible, we will use accurate, fair, and representative data sets to ensure these are inclusive. We will ensure that any AI decisions made are understandable and interpretable by all stakeholders. This could involve documenting the logic behind AI decisions, providing clear explanations, and making sure that the reasoning is accessible to non-technical users.

All reasonable efforts will be made to identify any bias within an AI system we use, to ensure any bias has either been eradicated or mitigated to the point where it is within an acceptable level of risk. We are open and transparent about any bias within an AI system (that we are aware of) and how we manage this.

Where AI is used to create content, there are appropriate checks and safeguards in place to ensure:

- There is appropriate content moderation by humans, to minimise the potential for errors and bias/defamatory phrases, etc.
 - We are open and transparent if and when content has been created solely by AI in response to specific prompts.
 - AI created content is either self-evident or clearly identified.
 - It will not be used for purposes where the use of AI has been specifically not permitted.
-

AI Environmental Considerations

FEAST is aware of the environmental impact of AI due to its very high energy and water consumption. We will take this into account when considering our environmental impact and seek to make use of any emerging technologies that will help to minimise or mitigate this.

AI Acceptable Use & Application

FEAST endeavours to:

- Understand that AI tools may be useful but are not a substitute for human judgment and creativity.
- Understand that many AI tools are prone to “hallucinations,” false answers or information, or information that is stale, and therefore responses must always be carefully verified by a human.
- Verify that any response from an AI tool that we intend to rely on, or use is accurate, appropriate, not biased, not a violation of any other individual or entity’s intellectual property or privacy, and consistent with FEAST policies and applicable laws. Furthermore, we understand that AI provided referencing can be incorrect.

FEAST will never:

- Use AI tools to make or help make employment decisions about applicants or employees, including recruitment, hiring, retention, promotions, transfers, performance monitoring, discipline, demotion, or terminations.
 - Upload or input any confidential, proprietary, or sensitive information into any AI tool.
 - Represent work generated by an AI tool as being our own original work.
-

AI Legal Compliance

FEAST will take all reasonable steps to identify copyrighted material. For any such material we use, we will ensure we have their copyright agreement, or it falls within 'fair use', or other exception to copyright, or the Open Government Licence (OGL), or some other free use category.

We will not knowingly use any online material, such as from social media accounts or online galleries, which has been marked as 'NoAI', 'NoImageAI', or similar.

We will take all reasonable steps to ensure that our use of AI does not have a negative impact on the legal rights and/or liberties of individuals or groups and complies with the Data Protection Act.

In particular, we will ensure that for any AI use of our data, the data is clean, complete, compliant and we have appropriate consent, particularly the safeguarding of sensitive personal information.

Cyber Security Policy in an AI enabled world

FEAST has robust cyber security procedures that staff and trustees are aware of and complies with consistently to minimise the risk of AI scams and disinformation.

Contacting Us

If you have any questions about this policy or our use of AI, please contact us at info@feastwithus.org.uk.

Reviewed and approved on 7 April 2025 by Caroline Monkhouse Flower, CEO, and Helen Burgess, Chair of the Board of Trustees, on behalf of the board.

Signed by:



Helen Burgess
Chair of the Board of Trustees

FEAST With Us, Registered Charity in England and Wales
Charity Number: 1172884
Address: 2A, The Quadrant, Epsom, KT17 4RH

Appendix 1 – Glossary of Terms

Algorithm

A set of instructions used to perform tasks (such as calculations and data analysis) usually using a computer or another smart device.

Algorithmic bias

AI systems can have bias embedded in them, which can manifest through various pathways including biased training datasets or biased decisions made by humans in the design of algorithms.

Artificial intelligence (AI)

The UK Government’s 2023 policy paper on ‘A pro-innovation approach to AI regulation’ defined AI, AI systems or AI technologies as “products and services that are ‘adaptable’ and ‘autonomous’.” The adaptability of AI refers to AI systems, after being trained, often developing the ability to perform new ways of finding patterns and connections in data that are not directly envisioned by their human programmers. The autonomy of AI refers to some AI systems that can make decisions without the intent or ongoing control of a human.

Artificial general intelligence

Sometimes known as general AI, strong AI or broad AI, this often refers to a theoretical form of AI that can achieve human-level or higher performance across most cognitive tasks. See also Superintelligence.

Artificial neural network

A computer structure inspired by the biological brain, consisting of a large set of interconnected computational units (‘neurons’) that are connected in layers. Data passes between these units as between neurons in a brain. Outputs of a previous layer are used as inputs for the next, and there can be hundreds of layers of units. An artificial neural network with more than 3 layers is considered a deep learning algorithm. Examples of artificial neural networks include Transformers or Generative adversarial networks.

Automated decision-making

A term that the Office for AI, within the Department for Science, Innovation and Technology, refers to in an Ethics, Transparency and Accountability Framework for Automated decision-making as “both solely automated decisions (no human judgement involved) and automated assisted decision-making (assisting human judgement).” AI systems are increasingly being used by the public and private sector for automated decision-making.

Compute

Compute is defined by the Independent Review of the Future of Compute as ‘the systems assembled at scale to tackle computational tasks beyond the capabilities of everyday computers. This includes both physical supercomputers and the use of cloud provision to tackle high computational loads.’ Compute is a driver of AI developments.

Computer vision

This focuses on programming computer systems to interpret and understand images, videos and other visual inputs and take actions or make recommendations based on that information.

Applications include object recognition, facial recognition, medical imaging analysis, navigation and video surveillance.

Deep learning

A subset of machine learning that uses artificial neural networks to recognise patterns in data and provide a suitable output, for example, a prediction. Deep learning is suitable for complex learning tasks, and has improved AI capabilities in tasks such as voice and image recognition, object detection and autonomous driving.

Deepfakes

Pictures and video that are deliberately altered to generate misinformation and disinformation.

Advances in generative AI have lowered the barrier for the production of deepfakes.

Disinformation

The UK Government defines disinformation as the “deliberate creation and spreading of false and/or manipulated information that is intended to deceive and mislead people, either for the purposes of causing harm, or for political, personal or financial gain”. Advances in generative AI have lowered the barrier for the production of disinformation, misinformation, and deepfakes.

Educational technology

Technologies specifically developed to facilitate teaching and learning which may or may not encompass AI.

Fine-tuning

Fine-tuning a model involves developers training it further on a specific set of data to improve its performance for a specific application.

Foundation models

A machine learning model trained on a vast amount of data so that it can easily be adapted for a wide range of general tasks, including being able to generate outputs (generative AI). See also large language models.

Frontier AI

Defined by the Government Office for Science as ‘highly capable general-purpose AI models that can perform a wide variety of tasks and match or exceed the capabilities present in today’s most advanced models’. Currently, this primarily encompasses a few large language models including

ChatGPT (OpenAI)

Claude (Anthropic)

and Gemini (Google)

Generative AI

An AI model that generates text, images, audio, video or other media in response to user prompts. It uses machine learning techniques to create new data that has similar characteristics to the data it was trained on. Generative AI applications include chatbots, photo and video filters, and virtual assistants.

General-purpose AI

Often refers to AI models that can be adapted to a wide range of applications (such as Foundation Models). See also narrow AI.

Generative adversarial networks

Generative adversarial networks are a driver of recent AI developments. These are made up of two sub artificial neural networks: a generator network and a discriminator network. The generator network is fed training data and generates artificial data based on patterns in training data. The discriminator network compares the artificially generated data with the ‘real’ training data and feeds back to the generator network where it has detected differences. The generator then alters its parameters. Over time the generator network learns to generate more realistic data, until the discriminator network cannot tell what is artificial and what is ‘real’ training data and the AI model generates the desired outcomes. See also artificial neural networks and transformers.

Graphical processing units

These are similar to central processing units, found on a typical home computer. Graphical processing units have been used since the 1970s in gaming applications and have been designed to accelerate computer graphics and image processing. In the past decade, graphical processing units have been increasingly applied in the training of large machine learning models after they were found to be effective in processing large amounts of data in parallel.

Hallucinations

Large language models, such as ChatGPT, are unable to identify if the phrases they generate make sense or are accurate. This can sometimes lead to inaccurate results, also known as ‘hallucination’ effects, where large language models generate plausible sounding but inaccurate text. Hallucinations can also result from biases in training datasets or the model’s lack of access to up-to-date information.

Interpretability

Some machine learning models, particularly those trained with deep learning, are so complex that it may be difficult or impossible to know how the model produced the output. Interpretability often describes the ability to present or explain a machine learning system’s decision-making process in terms that can be understood by humans. Interpretability is sometimes referred to as transparency or explainability.

Large language models

A type of foundation model that is trained on vast amounts of text to carry out natural language processing tasks. During training phases, large language models learn parameters from factors such

as the model size and training datasets. Parameters are then used by large language models to infer new content. Whilst there is no universally agreed figure for how large training datasets need to be, the biggest large language models (frontier AI) have been trained on billions or even trillions of bits of data. For example, the large language model underpinning ChatGPT 3.5 (released to the public in November 2022) was trained using 300 billion words obtained from internet text. See also natural language processing and foundation models.

Machine learning

A type of AI that allows a system to learn and improve from examples without all its instructions being explicitly programmed. Machine learning systems learn by finding patterns in training datasets. They then create a model (with algorithms) encompassing their findings. This model is then typically applied to new data to make predictions or provide other useful outputs, such as translating text. Training machine learning systems for specific applications can involve different forms of learning, such as supervised, unsupervised, semi-supervised and reinforcement learning.

Misinformation

The UK Government defines misinformation as “the inadvertent spread of false information”. Advances in generative AI have lowered the barrier for the production of disinformation, misinformation, and deepfakes.

Narrow AI

Sometimes known as weak AI, these AI models are designed to perform a specific task (such as speech recognition) and cannot be adapted to other tasks. See also general-purpose AI.

Natural language processing

This focuses on programming computer systems to understand and generate human speech and text. Algorithms look for linguistic patterns in how sentences and paragraphs are constructed and how words, context and structure work together to create meaning. Applications include speech-to-text converters, online tools that summarise text, chatbots, speech recognition and translations. See also large language models.

Open-source

Open-source often means the underlying code used to run AI models is freely available for testing, scrutiny and improvement.

Reinforcement learning

A way of training machine learning systems for a specific application. An AI system is trained by being rewarded for following certain ‘correct’ strategies and punished if it follows the ‘wrong’ strategies. After completing a task, the AI system receives feedback, which can sometimes be given by humans (known as ‘reinforcement learning from human feedback’). In the feedback, positive values are assigned to ‘correct’ strategies to encourage the AI system to use them, and negative values are assigned to ‘wrong’ strategies to discourage them, with the classification of ‘correct’ and ‘wrong’ depending on a pre-established outcome. This type of learning is useful for tweaking an AI model to follow certain ‘correct’ behaviours, such as fine-tuning a chatbot to output a preferred style,

tone or format of language. See also supervised learning, unsupervised learning and semi-supervised learning.

Responsible AI

Often refers to the practice of designing, developing, and deploying AI with certain values, such as being trustworthy, ethical, transparent, explainable, fair, robust and upholding privacy rights.

Robotics

Machines that are capable of automatically carrying out a series of actions and moving in the physical world. Modern robots contain algorithms that typically, but do not always, have some form of artificial intelligence. Applications include industrial robots used in manufacturing, medical robots for performing surgery, and self-navigating drones.

Semi-supervised learning

A way of training machine learning systems for a specific application. An AI system uses a mix of supervised and unsupervised learning and labelled and unlabelled data. This type of learning is useful when it is difficult to extract relevant features from data and when there are high volumes of complex data, such as identifying abnormalities in medical images, like potential tumours or other markers of diseases. See also supervised learning, unsupervised learning, reinforcement learning and training datasets.

Superintelligence

A theoretical form of AI that has intelligence greater than humans and exceeds their cognitive performance in most domains. See also artificial general intelligence.

Supervised learning

A way of training machine learning systems for a specific application. In a training phase, an AI system is fed labelled data. The system trains from the input data, and the resulting model is then tested to see if it can correctly apply labels to new unlabelled data (such as if it can correctly label unlabelled pictures of cats and dogs accordingly). This type of learning is useful when it is clear what is being searched for, such as identifying spam mail. See also semi-supervised learning, unsupervised learning, reinforcement learning and training datasets.

Training datasets

The set of data used to train an AI system. Training datasets can be labelled (for example, pictures of cats and dogs labelled 'cat' or 'dog' accordingly) or unlabelled.

Transformers

Transformers have greatly improved natural language processing, computer vision and robotic capabilities and the ability of AI models to generate text. A transformer can read vast amounts of text, spot patterns in how words and phrases relate to each other, and then make predictions about what word should come next. This ability to spot patterns in how words and phrases relate to each other is a key innovation, which has allowed AI models using transformer architectures to achieve a

greater level of comprehension than previously possible. See also artificial neural networks and generative adversarial networks.

Unsupervised learning

A way of training machine learning systems for a specific application. An AI system is fed large amounts of unlabelled data, in which it starts to recognise patterns of its own accord. This type of learning is useful when it is not clear what patterns are hidden in data, such as in online shopping basket recommendations (“customers who bought this item also bought the following items”). See also semi-supervised learning, supervised learning and reinforcement learning and training datasets.